Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

may, in one sense, have a range from zero to infinity; but, if our only measure is one half of infinity, we will, for all practical purposes, assign a value of two to the limiting velocity.

A still more interesting corollary of the hypothesis here presented is that the physicist may know when he has discovered a fundamental physical element by observing whether or not it exhibits a range of magnitudes characterized by a frequency distribution without a mode.

ON THE CORRELATION BETWEEN A VARIABLE AND THE DEVIATION OF AN ASSOCIATED BUT NOT DEPENDENT VARIABLE FROM ITS PROBABLE VALUE

Let r_{xy} be the correlation between two variables, x, y, which are so related to each other physically that y may be considered a dependent variable in the special sense in which this term has been used heretofore. Pearson and I have shown[1] that

$$r_{xz}=\frac{r_{xy}-V_x/V_y}{\sqrt{1-r_{xy}^2+(r_{xy}-V_x/V_y)^2}}$$

where V denotes the coefficient of variation of x and y, and z is to be read as the deviation of y from its probable value, that is

$$z=y-px$$
$$p=\bar{y}/\bar{x}.$$

Now let x and y be associated with a third variable, say w. The correlations r_{wx}, r_{wy}, r_{xy}, r_{xz}, and the partial correlation ${}_wr_{xy}$ may be readily determined. In certain cases the correlation r_{wz} and the partial correlation ${}_wr_{xz}$ may be of service.

$\bar{z}=0$; σ_z may be readily determined.[2] We require, therefore, merely the product moment for w and z. This might be determined by working directly with the individual values of the variable $z=y-px$. It is more convenient to proceed as follows:

Let $\Sigma(x)$, $\Sigma(y)$ denote the summation of the values of x and y for the n individuals of any class of w. Let S denote summation throughout the population or summation of the summations within the classes defined by the values of w.

Then within any class of n individuals defined by the value of w

$$\Sigma(z)=\Sigma(y)-p\Sigma(x).$$

For the whole series of $N=S(n)$ individuals

$$S(wz)=S[wS(z)]=S\left\{w[\Sigma(y)-p\Sigma(x)]\right\}.$$

---

[1] Harris, "The correlation between a variable and the deviation of a dependent variable from its probable value," *Biometrika*, 6 : 438–443. 1909.

[2] Harris, "Further illustrations of the applicability of a coefficient measuring the correlation between a variable and the deviation of a dependent variable from its probable value," *Genetics*, 3 : 328–352. 1918.

Finally, remembering that $\bar{z} = 0$ and expressing r in terms of product moments about 0 as origin.[1]

$$r_{wz} = \frac{S(wz)/N - \bar{w}\bar{z}}{\sigma_w \sigma_z}$$

and remembering that $\bar{z} = 0$, we may write

$$r_{wz} \frac{S\left\{ w[\Sigma(y) - p\Sigma(x)] \right\} / N}{\sigma_w \sigma_z}.$$

The means of arrays for testing the form of the regression curve of z on w are given by

$$[\Sigma(y) - p\Sigma(x)]/n.$$

J. ARTHUR HARRIS

CALCULATION OF THE CORRELATION RATIO

The correlation coefficient r is a good measure of correlation only when the regression is linear. It is necessary, therefore, before placing any reliance upon the computed r to examine the data for linearity of the regression lines. One common test of linearity, Blakeman's, compares the value of $(\eta^2 - r^2)$ with its probable error, η being a correlation ratio. In applying this test it is then necessary to calculate along with r the correlation ratios.

Recently the writer, with a number of assistants, among whom Mr. F. G. Wahlen should be mentioned, had occasion to compute several hundred correlation coefficients, each of which was tested by Blakeman's test. In the progress of the work, in which much use was made of the slide rule, there was developed the tabular form for the calculation of r with the two correlation ratios η_{yx} and η_{xy}, which is the content of this paper. By the addition of two columns to the tabular method of computing r used by Professor H. L. Rietz, the two correlation ratios are easily computed.

This method of calculating the correlation ratios in connection with the calculation of the correlation coefficient is shown in the accompanying numerical illustration. The various columns in connection with the correlation table explain themselves in the headings, with the exception of the columns headed $S_{s'}$ and S_s. The column headed $S_{s'}$ is computed as follows: Let s' be the number of any designated row of the correlation table, say the third row from the top. Then for that row each number is multiplied by the corresponding number in the column headed X. The algebraic sum of these products is $S_{s'}$. For example, for the third row from the top

$$\begin{array}{rl}
9 \cdot (-1) = & -9 \\
26 \cdot 0 = & 0 \\
7 \cdot 1 = & 7 \\
5 \cdot 2 = & \underline{10} \\
 & 8 = S_{s'} \text{ for } s' = 3.
\end{array}$$

The column headed S_s is computed in a similar way.

---

[1] Harris, "The arithmetic of the product moment method of calculating the coefficient of correlation." *Amer. Nat.*, 44 : 693–699. 1910.